FROM *BAREFOOT* TO *DAUBERT* TO *JOINER*: TRIPLE PLAY OR DOUBLE ERROR?

Michael H. Gottesman*

This paper traces the Supreme Court's erratic course over the past fifteen years in addressing the question whether federal judges may bar expert scientific testimony on the ground that the testimony is "unreliable." From the initial stance that no such reliability threshold exists (*Barefoot v. Estelle*),[1] the Court moved to the view that there is a reliable threshold—expert testimony will be admitted only when the expert's methodology conforms to the dictates of "good science" (*Daubert v. Merrell-Dow Pharmaceuticals, Inc.*),[2] and more recently moved again, this time to the view that district judges have broad discretion to admit or exclude expert testimony based upon their assessment of the reliability of both the methodology and the ultimate conclusions reached (*General Electric Co. v. Joiner*).[3] This journey has been marked from beginning to end with formalist opinions that do not address the *wisdom* of the choices being made. The opinions in *Barefoot* and *Joiner* announce their conflicting results simply as *ipse dixits*, while the *Daubert* opinion relies solely on dictionary definitions of two words ("scientific" and "knowledge") in Rule 702 of the Federal Rules of Evidence ("FRE").[4] Given that the controlling words in the Federal Rules of Evidence have remained unchanged throughout the period of these decisions, dictionaries seem an unsatisfying explanation for the divergent results.

This paper identifies the unarticulated reasons that likely explain this erratic journey. It also critiques the holdings in *Daubert* and *Joiner*, using two indices the Court did not consider: the intent of Congress when it enacted the FRE in 1975, and the public good. Finally, the paper speculates on the future of expert scientific testimony in light of the recent *Joiner* opinion.

---

* Professor of Law, Georgetown University Law Center. The author argued the *Daubert* and *Joiner* cases in the Supreme Court on behalf of the plaintiffs, and thanks Julie Ann Campbell for valuable research assistance in the preparation of this paper.

1. 463 U.S. 880 (1983).
2. 509 U.S. 579 (1993).
3. 118 S. Ct. 512 (1997).
4. FED. R. EVID. 702. For the text of Rule 702, see *infra*, note 17.

I. A CIRCUITOUS ODYSSEY: FROM *BAREFOOT* TO *DAUBERT* TO *JOINER*

A. Barefoot

In 1983, in *Barefoot v. Estelle*,[5] the Supreme Court considered whether the due process rights of a defendant in a state criminal trial are violated by the admission of psychiatric testimony regarding the defendant's future dangerousness. The American Psychiatric Association had filed an amicus brief stating that the mainstream of the psychiatric profession believed that psychiatrists' efforts to predict future dangerousness were totally unreliable.[6] The Court held that admitting the testimony was not a deprivation of due process, declaring that the safeguard against unreliable expert testimony is not exclusion of the evidence, but trust in the adversary system:

> If [the four psychiatrists whose testimony was at issue] are so obviously wrong and should be discredited, there should be no insuperable problem in doing so by calling members of the Association [to testify against them].... We are unconvinced, however, at least as of now, that the adversary process cannot be trusted to sort out the reliable from the unreliable evidence.[7]

In a footnote, the Court continued:

> All of these professional doubts about the usefulness of psychiatric predictions can be called to the attention of the jury.... Petitioner's entire argument...is founded on the premise that the jury will not be able to separate the wheat from the chaff. We do not share in this low evaluation of the adversary process.[8]

The court agreed with the district court's finding that the jury's purpose "is to sort out the true testimony from the false, the important matters from the unimportant matters, and, when called upon to do so, to give greater credence to one party's expert witnesses than another's. Such matters occur routinely in the American judicial system, both civil and criminal."[9]

That no constitutional violation occurs when a court trusts juries to "sort out the reliable from the unreliable" expert testimony is hardly remarkable. What *is* noteworthy is that the Court buttressed its conclusion that the practice is constitutional by noting that this stance is embraced by the FRE:

> [T]he rules of evidence generally extant at the federal...level[] anticipate that relevant, unprivileged evidence should be admitted

---

5. 463 U.S. 880 (1983).

6. Brief Amicus Curiae for the American Psychiatric Association, Barefoot v. Estelle, 463 U.S. 880 (1983) (No. 82-6080).

7. *Barefoot*, 463 U.S. at 900–01.

8 *Id.* at 901 n.7.

9. *Id.* at 902.

> and its weight left to the fact-finder, who would have the benefit of cross examination and contrary evidence by the opposing party.[10]

At the time the Supreme Court penned these words, no federal appeals court had ever ruled that judges could condition the admissibility of expert testimony in a civil case on a demonstration of its reliability.[11] Courts screened proposed expert testimony to assure that the expert was qualified and that the proffered testimony was relevant to the issues in the case, but not for reliability.[12] This was in accord with the views of Dean Charles McCormick, whose influential treatise, published before enactment of the FRE, disapproved screening of expert testimony for reliability.[13]

B. Daubert

A decade after *Barefoot*,[14] in *Daubert v. Merrell-Dow Pharmaceuticals, Inc.*,[15] the Supreme Court shifted course, declaring that "under the [FRE] the trial judge must ensure that any and all scientific testimony or evidence admitted is not only relevant, but reliable."[16] The Court found this requirement implicit in Rule 702's description of expert scientific testimony as testimony about "scientific...knowledge."[17] The Court reasoned:

---

10. *Id.* at 898.

11. The first such appellate decision appears to be *Barrel of Fun, Inc. v. State Farm Fire & Casualty Co.*, 739 F.2d 1028, 1031 (5th Cir. 1984). Before the FRE were enacted, at a time when evidence rules were judicially created, a number of lower courts had adopted the *Frye* rule, precluding the introduction in criminal cases of evidence derived from novel scientific techniques until "the thing from which the deduction is made [is] sufficiently established to have gained general acceptance in the particular field in which it belongs." Frye v. United States, 293 F. 1013, 1014 (D.C. Cir. 1923). Between enactment of the FRE in 1975 and the *Barrel of Fun* decision in 1984, a few courts continued to apply *Frye* in criminal cases, but prominent evidence scholars reported that "the [*Frye*] general acceptance test has been rejected by an increasing number of courts, and attacked by commentators, who have labeled the test 'infamous,' 'a sport,' 'archaic,' and 'antiquated on the day of its pronouncement.'" PAUL C. GIANELLI & EDWARD J. IMWINKELRIED, SCIENTIFIC EVIDENCE 13–14 (1986).

12. Additionally, courts had discretion to exclude any evidence, including expert testimony, despite its relevance, "if its probative value is substantially outweighed by the danger of unfair prejudice, confusion of the issues, or misleading the jury, or, by considerations of undue delay, waste of time, or needless presentation of cumulative evidence." FED. R. EVID. 403.

13. CHARLES T. MCCORMICK, HANDBOOK OF THE LAW OF EVIDENCE §170, at 363–64 (1954).

14. Between *Barefoot* and *Daubert* the Court issued one other opinion that appeared to reinforce *Barefoot's* perspective about the admissibility of expert testimony. In 1988, in *Beech Aircraft Corp. v. Rainey*, 488 U.S. 153, 169 (1988), the Court again declared that the FRE pursued a "general approach of relaxing the traditional barriers to 'opinion' testimony."

15. 509 U.S. 579 (1993).

16. *Id.* at 589.

17. Rule 702 provides, "[i]f scientific, technical, or other specialized knowledge

> The adjective "scientific" implies a grounding in the methods and procedures of science. Similarly, the word "knowledge" connotes more than subjective belief or unsupported speculation. The term "applies to any body of known facts or to any body of ideas inferred from such facts or accepted as truths on good grounds."... Of course, it would be unreasonable to conclude that the subject of scientific testimony must be "known" to a certainty; arguably, there are no certainties in science.... But, in order to qualify as "scientific knowledge," an inference or assertion must be derived by the scientific method.... In short, the requirement that an expert's testimony pertain to "scientific knowledge" establishes a standard of evidentiary reliability.[18]

As Professor Paul Giannelli has observed, *Daubert* cannot be squared with *Barefoot*.[19] What had happened in the intervening decade to cause this turnaround in the Court's position? One cannot tell from the face of the *Daubert* opinion. No change had occurred in the "federal rules" described by the *Barefoot* Court. The *Daubert* opinion does not even acknowledge *Barefoot*, much less attempt to rationalize its departure from the understanding of the FRE announced in *Barefoot*. The *Daubert* Court treated the issue as one of first impression and professed to find in the dictionary definitions of key words of FRE 702 a requirement of reliability.

In fact, a great deal had happened in that intervening decade, and it is not idle to speculate that the Court had these developments in mind when it shifted course in *Daubert*. Many federal appeals court judges had grown increasingly restive about what they perceived to be a phenomenon of venal experts saying anything the parties paying their fees wanted, without conformity to scientific discipline. A typical expression of this impatience appeared in a dissenting opinion of Seventh Circuit Judge Richard Posner:

> [The expert's testimony was] either [that] of a crank, or, what is more likely, of a man who is making a career out of testifying for plaintiffs in automobile accident cases in which a door may have opened; at the time of trial he was involved in 10 such cases. His testimony illustrates the age-old problem of expert witnesses who are "often the mere paid advocates or partisans of those who employ and pay them, as much so as the attorneys who conduct the suit. There is hardly anything, not palpably absurd on its face, that cannot now be proved by some so-called 'experts.'"[20]

---

will assist the trier of fact to understand the evidence or to determine a fact in issue, a witness qualified as an expert by knowledge, skill, experience, training, or education, may testify thereto in the form of an opinion or otherwise." FED. R. EVID. 702.

18. *Daubert*, 509 U.S. at 590 (citation omitted).

19. Paul C. Giannelli, Daubert*: Interpreting the Federal Rules of Evidence*, 15 CARDOZO L. REV. 1999, 2020–21 (1994).

20. Chaulk v. Volkswagen of Am., Inc., 808 F.2d 639, 644 (7th Cir. 1986) (Posner, J., dissenting) (quoting Keegan v. Minneapolis & St. Louis R.R., 78 N.W. 965, 966 (Minn. 1899)).

This same view was championed at book-length by Peter Huber in 1991, in *Galileo's Revenge: Junk Science in the Courtroom.* Huber did not simply rail against the practice. He found in it a threat to the American economy. The threat rested on two premises: (1) juries are not competent to recognize charlatans when they appear as experts in torts suits to testify that the defendant's product was defective and caused the plaintiff's injury, and (2) the result is the issuance of awards that deter manufacturers from introducing worthwhile products into the marketplace out of fear of unwarranted tort liability for injuries their products have not caused.[21]

Reflecting the restiveness within some quarters of the federal judiciary, in 1991 some judges on the Civil Rules Advisory Committee undertook an initiative to secure a change in the FRE to incorporate a reliability screen of the sort advocated by Huber. The proposed change would have amended Rule 702 by adding a requirement that expert testimony, to be admissible, must be "reasonably reliable."[22] The Committee proposing this change acknowledged that the FRE were "intended" to permit liberal introduction of expert testimony.[23] The Committee opined, however, that times had changed since 1975, when the FRE were adopted, and that "large expenditures for marginally useful expert testimony ha[ve] become commonplace," and "excesses...should be curtailed."[24] The Committee explained that its proposed "revision" (requiring that expert testimony be "reasonably reliable") would "limit the use, but increase the utility and reliability" of "testimony bearing on scientific and technical issues."[25] This proposal proved controversial and was formally withdrawn and then resubmitted when an Advisory

---

21. *See* PETER W. HUBER, GALILEO'S REVENGE: JUNK SCIENCE IN THE COURTROOM (1991). The evidence that Huber cites as supporting his thesis has been challenged at length in Kenneth J. Chesebro, *Galileo's Retort: Peter Huber's Junk Scholarship*, 42 AM. U. L. REV. 1637 (1993).

Huber also opined that juries are biased in favor of plaintiffs in torts suits because they see before them an injured person and a defendant with deep-pockets that can afford to pay for the injuries. As a result, juries may rule for plaintiffs even if they *do* recognize the speciousness of the opinion expressed by the plaintiffs' experts—in effect, they will ignore the judge's instructions and rule for the plaintiffs even though unconvinced that the defendants caused the plaintiffs' injuries. The only way to prevent this jury nullification, Huber argued, is to keep venal experts off the witness stand and thereby prevent the case from ever getting to a jury. Of course, this reasoning, if accurate, would justify a constitutional amendment to eliminate jury trials altogether in torts suits—we should not countenance a regime that we believe is systematically biased. After all, in every case in which the plaintiff has barely enough evidence to survive a summary judgment motion, the defendant will be at risk of a biased verdict no matter how persuasive its countering evidence. The exclusion of "unreliable" expert opinion, if done properly, will eliminate only those trials where the plaintiffs' case is the weakest—the cases where the defendants would be expected to be most successful even in jury trials.

22. *Preliminary Draft of Proposed Amendments to the Federal Rules of Civil Procedure and the Federal Rules of Evidence*, 137 F.R.D. 53, 73, 156 (1991) (proposal of Civil Rules Advisory Committee).

23. *Id.* at 156.

24. *Id.* at 156–57.

25. *Id.* at 156.

Committee on Evidence Rules was later created. The latter Committee had the proposal under advisement at the time *Daubert* was decided.

In 1992, a year before *Daubert*, the Bush Administration weighed in, lending its support to a proposed modification of the FRE to allow judges to monitor the reliability of expert testimony. A subcommittee of the President's Competitiveness Committee, chaired by Solicitor General Kenneth Starr, acknowledging that the FRE did not provide a judicial screen of the reliability of expert testimony, urged that the FRE be revised to incorporate such a screening role for federal judges.[26] The avowed purpose of the Competitiveness Committee's proposal was substantive: to reduce the number of torts judgments suffered by manufacturers, thereby enhancing their competitiveness.[27]

The FRE had been enacted by Congress as a statute (unlike the Federal Rules of Civil Procedure, which were promulgated by the Supreme Court and became effective in the absence of congressional action).[28] Accordingly, the FRE could not be changed to tighten admissibility standards without congressional action. If the FRE were as broad as *Barefoot* understood them, and as the judicial committees and Bush Administration acknowledged them to be, a reliability threshold for the admissibility of expert scientific testimony could not be added by judicial action alone.

With initiatives to amend the FRE to add a reliability screen underway, the Supreme Court in *Daubert* eliminated the need by finding a reliability screen *already* in FRE 702. The Court did so despite the fact that the language of Rule 702, its legislative history, and the context in which it was enacted all point strongly to the interpretation the Court had embraced in *Barefoot* a decade before: that the Rules enacted by Congress contained no threshold reliability requirement for the admissibility of expert testimony.[29]

The interpretation of Rule 702 of the FRE adopted in *Daubert* ostensibly did not empower the judge to assess whether the expert is testifying honestly. The Court did not suggest that it is the judge's function to decide whether the expert honestly believes in the opinion he or she is proffering. Rather, *Daubert* assumed that an expert might *really believe* the proffered testimony but authorized judges nonetheless to exclude the testimony as unreliable. This approach necessarily assumes that a judge is better equipped than an honestly-testifying expert to know whether the expert's opinion is reliable. That is an unlikely premise. To be sure, the judge will have heard the views of experts who disagree, but the testifying expert is just as aware of those competing views and, with far greater expertise than the

---

26. Dan Quayle, *Agenda for Civil Justice Reform in America*, 60 U. CIN. L. REV. 979, 999 (1992). The Committee stated that "[c]urrently, 'expert' witnesses are permitted to offer testimony even if their theories are unproven and not corroborated by other experts;" the proposed change would "exclud[e] fringe theories." *Id.*

27. *Id.* at 980.

28. Beech Aircraft Corp. v. Rainey, 488 U.S. 153, 163 (1988).

29. *See* Michael H. Gottesman, *Should Federal Evidence Rules Trump State Tort Policy? The Federalism Values* Daubert *Ignored*, 15 CARDOZO L. REV. 1837, 1854–63 (1994).

judge, has concluded that the contrary opinions are unconvincing. What is there to suggest that the judge is better able than the testifying expert to determine what is reliable? As two members of the *Daubert* Court observed, scientific knowledge is "far afield from the expertise of judges," and it is unlikely that Congress in the FRE intended judges "to become amateur scientists in order to perform [the reliability screening] role."[30]

If the enterprise launched in *Daubert* seems incongruous, one is left to speculate whether, perhaps, it was a proxy for another enterprise that went unstated. The words of *Barefoot* a decade before bear recalling: "We are unconvinced...*at least as of now*, that the adversary process cannot be trusted to sort out the reliable from the unreliable evidence...."[31] Perhaps in the ensuing decade the Court became convinced and came to share Peter Huber's concerns—many experts lie about what they really believe, juries are incapable of detecting when experts are lying, and, as a result, juries return verdicts requiring manufacturers to pay for injuries they did not cause. Consequently, manufacturers are deterred from marketing desirable products.

If this is what prompted the *Daubert* decision, it is not surprising that the Court concealed its reasons and resorted to formalistic dictionary-reading to explain its ruling. The Court could not admit that it was confronting the evil of venal experts by authorizing judges to discredit them as liars. The Seventh Amendment decrees a right to a jury in tort actions for damages tried in federal court, and it is doubtful that the Court could take from the jury the task of determining whether an expert witness is telling the truth as to what he or she believes. The judge's pre-trial determination of scientific "reliability"—an enterprise that professes to adjudge the scientific nature of the evidence rather than the credibility of the expert offering it—facilitates the exclusion of opinions the judge believes are not honestly held by the expert proffering them.[32]

It is a subject for fair debate whether judges as a class are better equipped than jurors to detect when experts are lying or to choose between competing opinions of even honestly testifying experts. The author's assessment of the

---

30. Daubert v. Merrell-Dow Pharm., Inc., 509 U.S. 579, 601 (1993), (Rehnquist, C.J., joined by Stevens, J., concurring in part and dissenting in part). By contrast, the majority professed "confiden[ce] that federal judges possess the capacity to undertake this review." *Id.* at 593. Even the majority did not suggest that federal judges have better capacity than qualified experts to undertake the review.

31. Barefoot v. Estelle, 463 U.S. 880, 901 (1983) (emphasis added).

32. Informal conversations between the author and judges at legal conferences about the capacity of trial judges to perform the task ostensibly assigned by *Daubert* support this suspicion. All judges who have deigned to discuss the matter with the author have acknowledged that the conundrum stated in text is unanswerable: the judge cannot presume to know better than an *honestly* testifying expert whether the expert's opinion is reliable. However, many judges challenged the assumption that all experts testify honestly. The author's untested hypothesis is that trial judges find the gatekeeper role assigned by *Daubert* a handy instrument for excluding expert testimony they believe to be dishonest, and that the Supreme Court fully intended this result.

arguments pro and con is briefly described in a footnote.[33] However, if one concludes that judges are better equipped, the conclusion would be true of many technical factual issues, not merely expert opinion.[34] Further, the framers of the Seventh Amendment did not qualify the right to a jury's determination of the relative competence of judges and juries to make accurate findings of fact.[35]

While the opinions for the Court in *Daubert* (and later in *Joiner*) contain not a whisper of the policy concerns expressed by Huber and the Bush

---

33. Judges have more education than the average juror and likely are better equipped to understand technical data and concepts. This is the strongest claim in support of judicial superiority. However, four factors weigh against the claim. First, only one trial judge sits on a case, in contrast to multiple jurors, and thus a greater risk exists that an aberrant notion will go uncorrected and determine the outcome. This risk is magnified by the holding in *Joiner*, described *infra* text accompanying notes 54–55. The holding in *Joiner* effectively insulates the trial judge's decision from serious appellate review. Second, judges usually make their *Daubert* determinations on the basis of affidavits submitted in support of and in opposition to summary judgment motions. This setting does not provide the same degree of illumination of the merits of the competing opinions (and of the credibility of the experts) as occurs through the process of direct and cross-examination at trial. Third, while it is possible for a trial judge to devote several days to reflecting upon the complex expert testimony in a single case, the demands of a busy docket are likely to preempt that luxury. The decision thus will be made without the time for extended reflection that a lengthy trial followed by jury deliberations will afford the jurors. Finally, trial judges have an incentive, however much they try to prevent its subconscious effect on their decisions, to clear their crowded dockets of cases that are likely to be time-consuming and, given the technicality of the evidence, tedious. A virtually unreviewable opportunity to shed cases that the judge thinks of doubtful merit must be a powerful temptation.

This analysis does not overlook Huber's concern that juries may be biased to rule in favor of injured persons at the expense of deep-pocket corporate defendants. As noted earlier, that charge if true would indict the civil jury system as a whole. Expert testimony should not be singled out as the only target to vindicate so sweeping an indictment. That we have constitutionalized the jury trial reflects a concern about a generic bias that may run the other way: that judges—overwhelmingly members of the intellectual and economic elite, and, in modern times, often graduates of a practice that consisted of representing corporations—on average are less empathetic than lay jurors to the plight of ordinary citizens suing powerful interests.

34. Another paper in this Symposium presents empirical evidence of the relative abilities of judges and juries to do fact-finding in complex cases, evidence that is far more favorable respecting jury capacity than Huber assumed (without systematic investigation) in advocating greater judicial control. Neil Vidmar, *The Performance of the American Civil Jury: An Empirical Perspective*, 40 ARIZ. L. REV. 849 (1998).

35. The Court's interest in limiting jury function on complex or important issues is also evidenced by two other recent decisions. *See* BMW of North America, Inc. v. Gore, 517 U.S. 559 (1996) (holding that due process forbids irrationally large amounts of punitive damages, thereby obliging both federal and state judges to review the size of jury awards of punitive damages); Markman v. Westview Instruments, Inc., 517 U.S. 370 (1996) (holding that, while patent infringement cases are entitled to jury trial under the Seventh Amendment, construction of the patent, including terms of art, may constitutionally be reserved to the judge, whose greater sophistication makes him or her better qualified for this task).

Administration, those concerns crept into Justice Breyer's concurring opinion in *Joiner*:

> [M]odern life, including good health as well as economic well-being, depends upon the use of artificial or manufactured substances, such as chemicals. And it may, therefore, prove particularly important to see that judges fulfill their *Daubert* gatekeeping function, so that they help assure that the powerful engine of tort liability, which can generate strong financial incentives to reduce, or to eliminate, production, points towards the right substances and does not destroy the wrong ones. It is, thus, essential in this science-related area that the courts administer the Federal Rules of Evidence in order to achieve the "end[s]" that the Rules themselves set forth, not only so that proceedings may be "justly determined," but also so "that the truth may be ascertained."[36]

Now, you might ask, what is the harm in this? Do we not want the truth to be ascertained so that, in Justice Breyer's words, we deter only "the right substances" and do not destroy the "wrong ones"? Of course we do. This is too glib, however, for there is an underlying question of policy that Justice Breyer's formulation ignores: how should we proceed when there is reason to suspect that a substance is "the right substance" for condemnation, but science has not arrived at a definitive judgment, one way or the other, about the causative properties of that substance? When the data is not determinative, is it more important to regulate the "right" substances, or more important to avoid regulating the "wrong" ones?

The degree of "reliability" that is imposed as a precondition to allowing the plaintiffs' experts to testify—and it is the plaintiffs' experts whose testimony is at stake, as the plaintiffs have the burden of proof and will be dismissed if admissible evidence is not proffered—is going to influence where, along the spectrum, the competing societal interests will be balanced. A government that deems the encouragement of new products more important than the risk of leaving victims uncompensated might insist upon a high degree of scientific certainty (or at least probability) before allowing a case to proceed. On the other hand, a

---

36. General Electric Co. v. Joiner, 118 S. Ct. 512, 520 (1997) (quoting FED. R. EVID. 102). While no similar discussion appears in the Court's opinions in either *Daubert* or *Joiner*, the Court in *Markman* did opine that judges are better equipped than juries to evaluate expert testimony about complex matters:

> It is, of course, true that credibility judgments have to be made about the experts who testify in patent cases, and in theory there could be a case in which a simple credibility judgment would suffice to choose between experts whose testimony was equally consistent with a patent's internal logic. But...in these cases a jury's capabilities to evaluate demeanor...are much less significant than a trained ability to evaluate the testimony in relation to the overall structure of the patent.... We accordingly think there is sufficient reason to treat construction of terms of art like many other responsibilities that we cede to a judge in the normal course of trial, notwithstanding its evidentiary underpinnings.

Markman, 517 U.S. at 389–90.

government that balances the policies differently and values compensation and deterrence over the societal benefits of risky substances might allow plaintiffs to recover on a showing that is less conclusive.

Presently, the choice among these policies rests with the states, or at least it is supposed to rest there. Tort law is state law, not federal law. However, when tort claims are litigated in federal court on the basis of diversity of citizenship jurisdiction, the federal courts, although obliged by *Erie* to honor a state's substantive tort law, may distort that law by imposing the *federal* admissibility rule of *Daubert* and dismissing suits that the state courts would have entertained. I have previously criticized *Daubert* because its reliability screen—which presumes that the question is one of scientific "fact" rather than a policy choice in the context of scientific uncertainty—is insensitive to state substantive law: the states' policy choices are ignored under the guise of administering a rule of evidence.[37] Professor Charles Nesson of Harvard cautioned long ago that insistence upon a high degree of scientific certainty of causation as a precondition to allowing a plaintiff to get to a jury in a tort case would elevate the plaintiff's burden far above the "likelier than not" standard that every state has chosen for resolving issues of causation.[38]

Daubert contemplates a uniform standard of "reliability" that will apply equally no matter what the issue being litigated in federal court. The burden of proof imposed upon the party proffering that evidence is irrelevant to the *Daubert* inquiry. Thus, a prosecutor in a criminal case who must prove her case beyond a reasonable doubt, a defendant in such a case who need only raise a reasonable doubt, and a litigant in a civil case who must prove the claim by a preponderance of the evidence, will confront the identical reliability threshold when they offer expert scientific testimony. However, these differences in burdens of proof reflect different choices about the willingness to tolerate outcomes that are predicated upon uncertainty about the underlying facts—differences that are grounded in varying policy considerations.[39]

The Court's insensitivity to these policy considerations has been made further manifest by its recent decision that a court can, consistent with the defendant's constitutional right to present relevant evidence in a criminal case, exclude the results of a polygraph examination proffered by the defendant that exonerates the defendant of the crime alleged.[40] The Court found the blanket exclusion of polygraph evidence in all military trials by Military Rules of Evidence, Rule 707 to be "a rational and proportional means of advancing the

---

37. Gottesman, *supra* note 29.

38. Charles Nesson, *Agent Orange Meets the Blue Bus: Factfinding at the Frontier of Knowledge*, 66 B.U. L. REV. 521, 527–31 (1986).

39. This point is made at length in Katherine Goldwasser, *Vindicating the Right to Trial by Jury and the Requirement of Proof Beyond a Reasonable Doubt: A Critique of the Conventional Wisdom About Excluding Defense Evidence*, 86 GEO. L.J. 621 (1998) (arguing that in light of the different burdens of proof in criminal cases defendants should be permitted to introduce evidence that would be excluded as insufficiently reliable if proffered by the prosecution).

40. United States v. Scheffer, 118 S. Ct. 1261 (1998).

legitimate interest in barring unreliable evidence."[41] Granted, the uncertain reliability of polygraph exams renders them relatively unpersuasive if they are offered to prove the defendant's guilt, which must be proved beyond a reasonable doubt. However, it does not logically follow that polygraph results suggesting the defendant's innocence are insufficient to raise in the minds of jurors a reasonable doubt about the defendant's guilt.[42]

In any event, wisely or not, *Daubert* created a federally-imposed solution to this state law question of how to deal with scientific uncertainty. *Daubert* did not, however, fix the *content* of that standard, and thus left unclear where, along the spectrum of uncertainty, the test would fall. The Court suggested some "factors" that "bear on the inquiry" of whether an expert's testimony is sufficiently reliable, but "we do not presume to set out a definitive checklist or test."[43] Rather, the Court invited lower courts to develop additional factors that might bear on "the reliability of evidence as ensured by the scientific validity of its underlying principles."[44] Some of the factors the Court did mention (error rate, testability, general acceptance)[45] smacked suspiciously of requiring a high level of certainty before evidence would be admitted, but the book was still open. After all, if factors such as these invariably were required, psychiatric testimony would never be admissible. Surely, the Court did not intend that!

Most observers viewed *Daubert* as the first step in the development of a principled and determinate jurisprudence governing the role of science in the courtroom, and to that end legions of lawyers and scholars penned tomes proffering their particular takes on how the enterprise should be perfected.[46] Alas, we learn with *Joiner*, that is not to be.

---

41. *Id.* at 1266.

42. Even Justice Stevens, the lone dissenter in *Scheffer*, was indifferent to the varying burdens of proof. He saw *other* reasons to distinguish a prosecutor's use from a defendant's—the prosecutor doesn't have a Sixth Amendment right, and "studies indicate that exculpatory polygraphs are more reliable than inculpatory ones." *Id.* at 1279 n.29. The difference in the parties' burdens of proof, however, went unmentioned.

43. Daubert v. Merrell-Dow Pharm., Inc., 509 U.S. 579, 593 (1993).

44. *Id.* at 595 n.12.

45. *Id.* at 593–94.

46. *See, e.g.*, Joe S. Cecil & Thomas E. Willging, *Accepting* Daubert's *Invitation: Defining a Role for Court-Appointed Experts in Assessing Scientific Validity*, 43 EMORY L.J. 995 (1994); Heidi Li Feldman, *Science and Uncertainty in Mass Exposure Litigation*, 74 TEX. L. REV. 1 (1995); Michael H. Gottesman, *Admissibility of Expert Testimony After* Daubert*: The "Prestige" Factor*, 43 EMORY L.J. 867 (1994); Jay P. Kesan, *An Autopsy of Scientific Evidence in a Post-*Daubert *World*, 84 GEO. L.J. 1985 (1996); Nancy A. Miller, Daubert *and Junk Science: Have Admissibility Standards Changed?*, 61 DEF. COUNS. J. 501 (1994); Note, *Improving Judicial Gatekeeping: Technical Advisors and Scientific Evidence*, 110 HARV. L. REV. 941 (1997); Jeffrey S. Parker, Daubert's *Debut: The Supreme Court, the Economics of Scientific Evidence, and the Adversarial System*, 4 SUP. CT. ECON. REV. 1 (1995); Laurens Walker & John Monahan, Daubert *and the Reference Manual: An Essay on the Future of Science in Law*, 82 VA. L. REV. 837 (1996).

C. Joiner *—First Impressions*

As it came to the Supreme Court, *Joiner* did not appear to be a case that would address the *content* of the *Daubert* reliability screen. The lone issue on which certiorari was sought and granted was: "[W]hat is the standard of appellate review for trial court decisions excluding expert testimony under *Daubert*...?"[47] In *Joiner*, the plaintiff[48] contended that PCBs manufactured by the defendants promoted (accelerated) the onset of his lung cancer.[49]

The defendants moved for summary judgment, contending that plaintiff could not show that PCBs cause or accelerate the onset of lung cancer. Thus, "general causation"—whether the challenged substance causes (or promotes) cancer in *anyone*—was placed in issue. Of course, even if that were proved, the question would remain whether it caused the *plaintiff's* injury ("specific causation"). Challenges to the reliability of expert testimony, however, most commonly arise in toxic torts cases at the general causation stage, as in *Joiner* (and *Daubert*). Once it is determined that the substance causes (or promotes) the type of injury the plaintiff is suffering, more finite techniques exist for assessing the probability that the substance caused (or promoted) the plaintiff's injury.[50]

---

47. Brief for Respondents at i, General Electric Co. v. Joiner, 118 S. Ct. 512 (1997) (No. 96-188).

48. There were two plaintiffs in *Joiner*, as Mrs. Joiner was suing for loss of consortium. For convenience, Mr. Joiner will be referred to as the plaintiff.

49. Plaintiff had been a cigarette smoker for eight years in his twenties, and contracted lung cancer at age 37 after working in an environment that exposed him (unknowingly) to massive doses of PCBs for nearly two decades. He did not dispute that smoking may have initiated the cell damage that led to his lung cancer, but it is unusual for anyone—even one who smoked far more than plaintiff—to suffer the actual onset of lung cancer at such an early age. Plaintiff contended that the onset of his cancer was accelerated by his exposure to PCBs. General Electric Co. v. Joiner, 118 S. Ct. 512, 516 (1997).

Joiner also contended that dioxins and furans that frequently are generated by the heating of PCBs also contributed to the acceleration of his lung cancer. The district court ruled that plaintiff had failed to identify evidence that would permit a jury to find that he had been exposed to dioxins or furans and granted summary judgment on these claims without addressing the admissibility of plaintiff's expert testimony that furans and dioxins promote lung cancer. The court of appeals reversed this ruling, holding that the record contained evidence of exposure to furans and dioxins. Defendants did not seek Supreme Court review of that ruling, and thus the claims involving those substances remain "open questions" to be litigated further in the district court. *Id.* at 519. As Justice Stevens explained more fully in his separate opinion concurring in part and dissenting in part:

> Petitioners do not challenge the Court of Appeals' straightforward review of the District Court's summary judgment ruling on exposure to furans and dioxins. As today's opinion indicates, *ante*, at 519, it remains an open question on remand whether the District Court should admit expert testimony that PCBs, furans and dioxins *together* promoted Joiner's cancer.

Id. at 522 n.3 (emphasis added).

50. For example, if the evidence on general causation shows that the substance more than doubles the incidence of the particular type of injury among those exposed to it, that alone will suggest that anyone exposed to the substance likelier than not contracted the

In *Joiner*, the plaintiffs countered the summary judgment motion with affidavits of two distinguished scientists, both of whom opined that likelier than not PCBs generally promote lung cancer and, specifically, likely accelerated the onset of Mr. Joiner's lung cancer.[51]

In *Joiner*, the district court granted the defendants' motion for summary judgment, ruling that plaintiffs' only evidence that PCBs promote lung cancer—the testimony of expert witnesses—was inadmissible under *Daubert's* reliability test. On appeal, the Eleventh Circuit reversed, applying an abuse of discretion standard of review embellished in the following respect: "Because the Federal Rules of Evidence governing expert testimony display a preference for admissibility, we apply a particularly stringent standard of review to the trial judge's exclusion of expert testimony."[52] Defendants' petition for certiorari challenged this gloss on the abuse of discretion standard. They asked the Supreme Court to hold that a single, non-stringent standard of abuse of discretion governs the appellate review of all rulings on the admissibility of evidence. Defendants requested that if the Court so

---

injury from it. 2 AMERICAN LAW INST., REPORTER'S STUDY, ENTERPRISE RESPONSIBILITY FOR PERSONAL INJURY 369–75 (1991). Even without that, the plaintiff's individual profile may make him or her *less likely* than the average person to have contracted the injury from *other* known causes—for instance, a non-smoker contracting lung cancer or a person contracting heart disease who does not have a family history of heart disease or any of the known indicators of increased likelihood (high cholesterol, obesity, etc.). Finally, states can address the issue of specific causation by allowing partial recoveries. If a substance is known to cause one-third of the injuries suffered by the population as a whole, but specific evidence respecting the effect of the substance on the plaintiff is unavailable, a state might opt to award the plaintiff one-third of his or her damages. *Id.*

51. Dr. Daniel Teitelbaum was a co-founder of both the American Academy of Clinical Toxicology and the American Board of Medical Toxicology. He has published more than forty articles on occupational and environmental toxicology and the epidemiology of toxic diseases in peer-reviewed journals. He has assisted a number of federal agencies in assessing the toxicity of chemicals and other substances—serving as chairman of an FDA panel, as a member of several EPA committees or panels, as a special consultant to OSHA on many occasions, and as a member of a blue-ribbon presidential panel that was formed to chart strategy for future EPA and NIH research concerning the health risks of emerging technologies. He has served as an advisor or panel member of several World Health Organization working groups. He has taught federal judges on the subjects of toxicology and epidemiology, under the auspices of the Federal Judicial Center. He has consulted on occupational and toxicology issues for a number of major corporations, including IBM, W.R. Grace, Amoco, Xerox, Motorola and Intel. Joint Appendix at 447–48, 473, 476–83, General Electric Co. v. Joiner, 118 S. Ct. 512 (1997) (No. 96-188).

Dr. Arnold Schecter is a tenured medical professor, and one of the few American physicians who works full time doing research on the toxic effects of PCBs, furans and dioxins. He has published over 100 peer-reviewed papers on these substances, and has served as a consultant respecting these substances to the EPA, U.S. Public Health Service, National Academy of Sciences, and World Health Organization. *Id.* at 399–401, 414–15, 424–42.

52. Joiner v. General Electric Co., 78 F.3d 524, 529 (11th Cir. 1996).

held, it remand the case to the court of appeals for application of the correct standard of review.[53]

The Supreme Court reversed on the standard of review, holding that "the question of admissibility of expert testimony...is reviewable under the abuse of discretion standard" and rejecting the court of appeals' view that rulings excluding expert testimony should be "subjected to a more searching standard of review."[54]

Had the Court stopped there and remanded the case as defendants had requested, its decision would have been uncontroversial. However, although it had already answered the only question presented, the Court went on to hand the defendants a victory greater than they had sought: "We hold that abuse of discretion is the appropriate standard. *We apply this standard and conclude that the District Court in this case did not abuse its discretion when it excluded certain proffered expert testimony.*"[55] The explanation provided by the Court for this second holding represents a marked amendment to the *content* of the reliability standard announced in *Daubert.*

To appreciate the significance of the route the Court traveled in the uninvited portion of its opinion, one must be aware of a fundamental debate within the scientific community about the methodology for evaluating whether a substance causes particular diseases.

D. The Debate about the Scientific Method

Although the Court in *Daubert* described science as if it were a monolithic slow march in search of enduring truth, in practice scientists are embarked on a variety of quite different missions. Scientists employ quite different methodologies in determining issues of causation, depending upon the reason they are undertaking the inquiry.

1. Theoretical Science's Methodology

In one corner is the theoretical scientist, whose mission is to search for enduring truths about the state of nature. Asked whether a particular chemical is carcinogenic, this scientist undertakes a variety of pertinent studies, gathers data, and assesses whether the data has become so definitive that a reliable answer can be declared. This scientist will not be satisfied by accumulating evidence that suggests a likelihood that a cause-and-effect relationship exists. Such evidence is exciting—it whets the appetite for further research (and may convince others to provide more money in pursuit of the inquiry)—but there is no rush: the search will go on until the evidence is so overwhelming (for or against causation) that announcement of a definitive conclusion is warranted.

---

53. Petition for Writ of Certiorari at 14 n.4, General Electric Co. v. Joiner, 118 S. Ct. 512 (1997) (No. 96-188).

54. *Joiner*, 118 S. Ct. at 517 (1997).

55. *Id.* at 515 (emphasis added). *See also id.* at 517–19.

The theoretical scientist will be interested in studies that show that animals contract cancer from the chemical but will not find this conclusive, as animal studies, while often predictive of human causation, are not always so. For like reasons, the fact that a chemical has a composition that resembles other chemicals known to be human carcinogens will not be regarded as definitive. Only evidence that links exposure to cancer in humans—for instance, human epidemiological evidence—is going to permit the theoretical scientist to arrive at closure. If people who are exposed to the chemical contract cancer at rates that are significantly higher than those who are not so exposed—sufficiently higher to negate the chance that the higher rates are mere random deviations from the norm—that evidence will be regarded as conclusive.[56]

Here is the rub. It is rare that such conclusive human epidemiological evidence is obtainable. Several reasons account for this. Consider the typical chemical that is introduced into the marketplace and then, after a period of time, gives rise to suspicion that it may be causing toxic harms. At the moment manufacturers discover this, they are likely to take steps to prevent a continuation of the harm: for both humanitarian and legal reasons, they will institute protective procedures for future use of the product or, if that is not possible, may withdraw the product from the market. In either event, human exposure (in unprotected form) will stop at the very point when it becomes relevant to ask whether it had caused injury to humans before that potential was noticed. While it may be possible to look backward and to inquire whether there was an elevated incidence among those exposed prior to this change, the absence of a monitored study will make evaluation of the evidence difficult if not impossible. (The reasons will become evident shortly.)

Bendectin, the product at issue in *Daubert* (an anti-nausea product taken by pregnant women) had been withdrawn from the American market shortly after the earliest reports that mothers who took it during pregnancy were giving birth to children with limb defects. PCBs, the product in *Joiner*, had been routinely used as a heat-retardant in electrical transformers, and it is in that context that Joiner, an electrician who repaired transformers for the utility company, suffered his extensive exposure. As reports began to emanate linking at least *some* cancers to exposure to PCBs, their use in transformers was discontinued. There simply were no human epidemiological studies that conclusively answered the questions, one way or the other, whether links existed between the substances at issue in *Daubert* and *Joiner* and the injuries the plaintiffs suffered and no way to conduct such studies prospectively.

In each instance, there were retrospective studies—examinations after the fact of populations that had been exposed to these substances before their dangerousness had been recognized—and those studies indicated elevated incidence of fetal limb deformation (in the case of Bendectin) and of lung cancer

56. Of course, this inquiry must be conducted separately for each type of cancer. Thus, the theoretical scientist will not be satisfied that a chemical causes lung cancer until there is conclusive epidemiological evidence that persons exposed to the chemical contract lung cancer at statistically significantly higher rates than those who are not.

(in the case of PCBs) among those exposed to these substances. However, the data was insufficient to permit a finding of "statistical significance" to assure that the elevations were not simply reflections of the routine randomness of natural phenomena.[57]

Moreover, human epidemiological studies are very difficult to conduct because, unlike with animal studies, scientists cannot control with exactitude the environment of humans. Humans are likely to move about and expose themselves to *other* potential carcinogens (for example, by smoking), which will confound the evidence the scientists are trying to examine. If the "control group" (those not exposed to PCBs, for example) are exposing themselves to other carcinogens that may cause lung cancer (carcinogens to which the PCB group is not exposed in equal measure), the control group will show up with the same level of lung cancer, and the causative effect of the PCBs will be masked. Of course, in prospective studies, scientists will try to account for this by excluding from the two groups people who have other exposures, but this assumes that scientists know what all the other carcinogens are and that the human subjects accurately report their exposures. What is more, humans are unlikely to be monitored in a way that will permit a determination of exactly how much exposure they have had to the suspected carcinogen and to other possible carcinogens.

These problems are magnified when scientists attempt to study a causation issue retrospectively, for there will have been no effort to segregate the groups that are to be compared and control their other exposures. In addition, it is impossible to get accurate data with hindsight about other exposures simply by interviewing all of the subjects (including, most importantly, those who have died following exposure).

These and other problems have led the Environmental Protection Agency to conclude, in published regulations, that human epidemiological studies "are useful in generating hypotheses and providing supportive data, *but can rarely be used to make a causal inference.*"[58] In consequence, a judicial stance insisting upon

---

57. Statistical significance is a conventional test used by theoretical scientists to identify correlations that deviate *so far* from the norm that it is extremely unlikely that they are mere random results. For example, if one flipped a coin 100 times, and then another 100 times, and then again and again, one would not expect to achieve 50 heads and 50 tails in each batch of 100 flips even if the coin were perfectly balanced. A certain amount of deviation from the norm—usually distributed in a bell-shaped curve around the norm—is expected. But if one kept getting results that were significantly skewed in one direction, for example, 90–10 heads, then 77–23 heads, then 86–14 heads, etc., one would soon be convinced that these are not the results of a well-balanced coin, but, rather, strong evidence that the coin is imbalanced in a way that causes it to land more often on its head.

The convention among theoretical scientists is to regard as inconclusive any results that are not *so far* from the norm as to occur less than five percent of the time by chance; or, put the other way, to treat only those results at the far ends of the bell curve as significant. That choice is made because theoretical scientists want enduring answers, not momentary best guesses, and thus want a high level of confidence that what they have observed is not a random deviation.

58. EPA Final Guidelines for Carcinogen Risk Assessment, 51 Fed. Reg. 33,992,

proof that would satisfy a theoretical scientist as to the "truth" of a claim of causation before regarding testimony of causation to be "reliable" would rarely allow plaintiffs to introduce evidence of causation. Because evidence of causation cannot be presented other than through experts, suits would be precluded in the vast range of cases where causation is suspected based on the available data.

It is important to stress that this does not mean that the evidence points to a lack of causation. Quite the contrary: in both *Daubert* and *Joiner*, the available data—human epidemiological data and otherwise—*is* suggestive of causation. Those humans known to be exposed to the challenged substances suffered the claimed effects in disproportionately high numbers (compared to those not known to be so exposed), and the available animal and chemical data also suggested a link. However, the human epidemiological results were not *so far* from the norm as to allow a finding of statistical significance at the conventional ninety-five percent level. To put it another way, if the defendants were called upon to prove a lack of causation, their efforts would fall even further from the mark. But, of course, the burden of producing admissible evidence rests with the party who has the burden of proof, and in torts cases the burden of proving causation is always on the plaintiff.

2. Practical Science's Methodology: Weight-of-the-evidence Approach

While theoretical scientists will not be satisfied, the question about which they are not satisfied is: "does substance *A* cause disease *B*?" These scientists are not in the "likelier than not" business. Other scientists, however, are in that business. Their mission is to make predictions about how likely it is that substance *A* is causing disease *B*. These include the scientists who advise environmental and public health regulators, as well as medical doctors who must make diagnoses in order to treat their patients.

Now, to be sure, these scientists are not solely interested in whether causation is more or less than fifty percent likely. Products might be banned, or safeguards taken, even if the likelihood of causation is less than fifty percent if the magnitude of the disease warrants such precautions. It is a mistake, however, to think that these scientists are simply asking whether it is possible (however unlikely) that substance *A* is causing disease *B*. For example, regulators will want to make a cost-benefit calculus before deciding to ban or limit a substance. One ingredient of that calculus is an assessment of how likely the feared effects may be.

The need being present, a well-developed set of scientific protocols now exists, operative when science does not have a conclusive answer, to assess how likely it is that a product causes specified diseases. Federal agencies have published these protocols—colloquially known as the "weight-of-the-evidence" methodology—in regulations. For example, the current EPA guidelines on the subject state:

> Evidence of possible carcinogenicity in humans comes primarily from two sources: long-term animal tests and epidemiological

33,995 (1986) (emphasis added).

> investigations. Results from these studies are supplemented with available information from short-term tests, pharmacokinetic studies, comparative metabolism studies, structure-activity relationships, and other relevant toxicological studies. *The question of how likely an agent is to be a human carcinogen should be answered in the framework of a weight-of-the-evidence judgment.* Judgments about the weight of the evidence involve considerations of the quality and adequacy of the data and the kinds and consistency of responses induced by a suspect carcinogen. There are three major steps to characterizing the weight of evidence for carcinogenicity in humans: (1) characterization of the evidence from human studies and from animal studies individually, (2) combination of the characterizations of these two types of data into an indication of the overall weight of evidence for human carcinogenicity, and (3) evaluation of all supporting information to determine if the overall weight of evidence should be modified.[59]

As is evident from this formulation (and from the more detailed portions of the regulation that elaborate upon the formulation), the answer is not derived linearly from some mathematical equation. Rather, scientists assemble the available data and then exercise their intellects to derive a probability assessment from the admittedly inconclusive data. The regulation calls for "evaluation" of the data to arrive at a "judgment" as to "how likely an agent is to be a human carcinogen." A group of distinguished scientists, led by Stephen J. Gould, filed an *amicus* brief in *Daubert*, explaining that this process requires dependence upon "subjective judgments."[60] These scientists quoted a basic text for the following proposition:

> In the end, a quality which lawyers should understand better than any—judiciousness—matters more than any. Scientists use both deductive and inductive inference to sustain the momentum of a continuing process of research.... The courts of law, and the courts of application, use inference to reach decisions about what action to take. Those decisions often cannot rest on certitudes.[61]

E. The Debate over Competing Methodologies in Toxic Torts Cases

This dichotomy of methodology between these two groups of scientists replicates itself on a regular basis in toxic torts cases. In both *Daubert* and *Joiner*, as, indeed, in a myriad of such cases, the defendants moved for summary judgment, filing one or more affidavits of scientists who opined that the defendant's substance has not been shown to cause the type of disease the plaintiff suffered. These affidavits recited that no substantial body of human epidemiological evidence shows with statistical significance any such causation, and absent such evidence, it

---

59. *Id.* at 33,996 (emphasis added).

60. Brief *Amici Curiae* of Physicians, Scientists and Historians of Science in Support of Petitioners at 8–9, *Daubert v. Merrell-Dow Pharmaceuticals, Inc.*, 509 U.S. 579 (1993) (No. 93–102).

61. *Id.* at 9–10 (quoting Susser, *Rules of Inference in Epidemiology*, REG. TOXICOLOGY & PHARMACOLOGY 116, 128 (1986)).

is impossible to establish causation. The defendants' affidavits demonstrate that there is insufficient evidence of causation to satisfy a theoretical scientist.

However, the plaintiffs countered with affidavits by multiple distinguished scientists,[62] stating that they had examined all available evidence under the methodology prescribed in the EPA guidelines, and that, on the basis of such evidence, it appears likelier than not that the substance causes this type of disease. The plaintiffs' affidavits demonstrate there is sufficient evidence of causation to satisfy a practical scientist.

In *Daubert*, the Supreme Court was explicit about what, particularly, had to be "reliable" for the expert testimony to be admissible: "[The] overarching subject [of the reliability inquiry] is the scientific validity—and thus the evidentiary relevance and reliability—of the *principles that underlie* a proposed submission. *The focus, of course, must be solely on principles and methodology, not on the conclusions that they generate.*"[63] This followed from the Court's approach to interpreting the words "scientific...knowledge" in FRE 702, which recognized that "it would be unreasonable to conclude that the subject of scientific testimony must be 'known' to a certainty; arguably, there are no certainties in science.... But, in order to qualify as 'scientific knowledge,' an inference must be *derived by the scientific method.*"[64]

The sixty-four dollar question after *Daubert* was whether the weight-of-the-evidence approach is a "scientifically valid" methodology for determining the issue of causation in torts cases. If it is, and if the testifying expert followed this methodology, *Daubert* appeared to say that the reliability threshold had been satisfied, and the expert's opinion respecting causation would be admissible. Other experts might quarrel with the *conclusions* reached by the expert applying this methodology, but disputes about conclusions were for the jury to decide and did not affect admissibility.

However, in its subsequent opinion in *Joiner*, addressing whether the district court had abused its discretion in excluding the testimony of the plaintiff's experts, the Supreme Court entirely ignored the framework for determining reliability it had constructed in *Daubert*. The Court nowhere mentioned the "factors" that *Daubert* had declared were to be used to determine whether an expert has followed the scientific method. Instead, the Court in *Joiner* proffered an essay granting federal district courts broad latitude in ruling on the admissibility of expert testimony without fear of appellate reversal.

It is possible to read the Court's opinion, as Justice Stevens does in his partial dissent, to hold that the district court has discretion whether to treat "weight

---

62. In *Daubert*, the plaintiff offered the testimony of eight experts. The Supreme Court stated that "each" of the eight "possessed impressive credentials." Daubert v. Merrell-Dow Pharm., Inc., 509 U.S. 579, 583 (1993). *See also id.* at 583 n.2, (describing the credentials of three of these experts and declaring that "[t]he credentials of the others are similarly impressive"). In *Joiner*, the plaintiff offered the testimony of two experts, whose impressive credentials are detailed *supra* note 51.

63. *Daubert*, 509 U.S. at 594–95 (emphasis added).

64. *Id.* at 590 (emphasis added).

of the evidence" as a reliable methodology. It is also possible to read the opinion as implicitly acknowledging the propriety of that methodology but empowering district courts to exclude evidence whenever they disagree with the inductive reasoning by which the expert employing that methodology arrived at his or her conclusion about the probability of causation.

To make sense of the Court's discussion, we must understand the reasoning of the lower courts in *Joiner*. The affidavits of the plaintiff's experts relied on the totality of available animal and epidemiological evidence respecting PCBs. These experts acknowledged that no single piece of evidence, standing alone, would permit a confident determination that causation was probable. They explained, however, that they had applied the EPA weight-of-the-evidence approach faithfully and concluded, upon consideration of all the available data, that it is likelier than not that PCBs promote (accelerate the onset of) lung cancer.

The trial judge, however, ignored the weight-of-the-evidence approach. She evaluated the items of data upon which the plaintiff's experts relied, one-by-one, and concluded that each, standing alone, did not establish causation. On that basis, she ruled the experts' conclusions unreliable, and thus inadmissible.

The court of appeals reversed the trial judge precisely because, the appellate court reasoned, it was legal error to disregard the weight-of-the-evidence approach. Justice Stevens aptly captured the conflict between the positions of the two lower courts, and allied himself with the court of appeals, in his partial dissent in *Joiner*:

> Unlike the District Court, the Court of Appeals expressly decided that a "weight of the evidence" methodology was scientifically acceptable. To this extent, the Court of Appeals' opinion is persuasive. It is not intrinsically "unscientific" for experienced professionals to arrive at a conclusion by weighing all available scientific evidence—this is not the sort of "junk science" with which *Daubert* was concerned. After all, as Joiner points out, the Environmental Protection Agency (EPA) uses the same methodology to assess risks, albeit using a somewhat different threshold than that required in a trial.[65]

The Supreme Court majority did not expressly declare whether weight of the evidence is a permissible methodology. Instead, it tracked the district court's reasoning, declaring at each step that it was not an abuse of discretion to reason as the district court had.

It first ruled that the district court had not abused its discretion in rejecting the experts' reliance on the results of animal studies showing that infant mice injected with high doses of PCBs contracted lung cancer at statistically significant levels. Joiner had received a lower dose, and these were *infant* mice: there was no study showing that "*adult* mice developed cancer after being exposed to PCBs."[66]

---

65. General Electric Co. v. Joiner, 118 S. Ct. 512, 522–23 (1997) (footnotes omitted).

66. *Id.* at 518 (emphasis added).

The Court did not explain what principle of science made results respecting adult mice more persuasive than infant mice, nor did it actually concede that results respecting adult mice would have been entitled to different treatment:

> Of course, whether animal studies can ever be a proper foundation for an expert's opinion was not the issue. The issue was whether *these* expert's opinions were sufficiently supported by the animal studies on which they purported to rely. The studies were so dissimilar to the facts presented in this litigation that it was not an abuse of discretion for the District Court to have rejected the experts' reliance on them.[67]

Notice that the Court in this passage has held non-abusive the district court's total disregard of the results of animal studies—they were deserving of no consideration—because they were unpersuasive *standing alone* to support a finding of causation. This tolerates the very approach that the court of appeals had declared legal error, and supports Justice Stevens' reading that the Supreme Court was allowing district courts to disapprove the weight-of-the-evidence approach.

The Supreme Court then turned to the several epidemiological studies, relied upon by plaintiff's experts, that had found elevated levels of lung cancer in humans exposed to PCBs. It held that the district court did not abuse its discretion by concluding that the first such study did not support the experts' conclusion because the authors of that study "were unwilling to say that PCB exposure had caused cancer among the workers they examined."[68] Of course, as this was the first such study, the authors were saying only that this study standing alone would not support such a conclusion, yet the Court deemed that sufficient to sweep it off the table.

The district court did not abuse its discretion in rejecting reliance on a second study, because the increased incidence of lung cancer "was not statistically significant."[69] It could permissibly ignore yet another study, which *did* show a statistically significant connection between PCBs and lung cancer, because the subjects of the study had also been subjected to other potential carcinogens.[70]

Justice Stevens, in his partial dissent, "fully agree[d]...[that] each of the studies on which the experts relied was *by itself* unpersuasive."[71] However, he reasoned, "a critical question remains unanswered."[72] Given that the experts did not purport to be basing their conclusions on any one study standing alone, but, rather, on the weight of their combined evidence, and given that this methodology is the one applied by the federal government, "why are their opinions inadmissible?"[73]

---

67. *Id.* (emphasis in original).
68. *Id.*
69. *Id.* at 519.
70. *Id.*
71. *Id.* at 523 (emphasis added) (Stevens, J., dissenting).
72. *Id.*
73. *Id.*

The majority did not respond to this in terms, nor did it acknowledge that this was the aspect of the district court's reasoning that had led the court of appeals to reverse. Nor did the majority explain why the evidence considered as a whole could be considered unpersuasive. These features of the opinion lend credence to Justice Stevens' observation that the Court has held that district courts are free to credit or discredit the weight-of-the-evidence methodology and are entitled to affirmance whichever choice they make.

The Court did, however, add a sentence at the very end of its opinion, in a passage that purports merely to be recapitulating its holding, that suggests a second possible reading of its opinion:

> We hold, therefore, that abuse of discretion is the proper standard by which to review a district court's decision to admit or exclude scientific evidence. We further hold that, because it was within the District Court's discretion to conclude that the studies upon which the experts relied were not sufficient, *whether individually or in combination*, to support their conclusions that Joiner's exposure to PCBs contributed to his cancer, the District Court did not abuse its discretion in excluding their testimony.[74]

This is the only hint that the majority might have thought the district court was obliged to credit the weight-of-the-evidence methodology as "scientific." However, if the Court did think the district court was so obliged, its opinion marks a departure from *Daubert* that is equally important.

If the experts here pursued a scientific methodology, and the district court's indictment was that the evidence upon which they relied was not sufficient "to support their conclusions," then the district court committed the very error that *Daubert* had warned against. To recall the words of that earlier opinion: "The focus, of course, must be solely on principles and methodology, not on the conclusions that they generate."[75] That the *Joiner* Court meant to scuttle *Daubert's* approach was not left to inference.

> Respondent points to *Daubert*'s language that the "focus, of course, must be solely on principles and methodology, not on the conclusions they generate."...He claims that because the District Court's disagreement was with the conclusion that the experts drew from the studies, the District Court committed legal error and was properly reversed by the Court of Appeals. But conclusions and methodology are not entirely distinct from one another. Trained experts commonly extrapolate from existing data. But nothing in either *Daubert* or the Federal Rules of Evidence requires a district court to admit opinion evidence which is connected to existing data only by the *ipse dixit* of the expert. A court may conclude that there is simply too great an analytical gap between the data and the

74. *Id.* at 519 (emphasis added).
75. Daubert v. Merrell-Dow Pharm., Inc., 509 U.S. 579, 595 (1993).

> opinion proffered.... That is what the District Court did here, and we hold that it did not abuse its discretion in so doing.[76]

Ipse dixit indeed! One might use that term to describe the Court's passage that begins "But nothing in...*Daubert*...requires..." Or so, at least, Justice Stevens thought, declaring the passage not "accurate."[77]

II. A CRITIQUE OF THE PRESENT STATE OF AFFAIRS

I stated earlier that *Daubert* was a wrong turn because it presumed that judges could know better than scientists whether the scientists' opinions are reliable, and because it was insensitive to the substantive policy choices that the state makes in determining what level of confidence is required to affix liability in a world of scientific uncertainty. *Joiner* has fundamentally changed the Court's stance, substituting one set of problems for another. In contrast to *Daubert*, which appeared to be launching the federal courts on a search for the determinative "scientific method," *Joiner* is a paean to "anything goes." Indeed, that the Court reached out to decide an issue that was not even presented is testimony to how anxious it was to send that message.

However one reads the majority in *Joiner*—whether as saying that district courts have free rein to approve or disapprove the weight-of-the-evidence methodology or as saying that they must approve that methodology but have a wide berth to agree or disagree with the conclusions reached by the experts applying that methodology—the decision represents a silly turn. Either version places too much discretion in the hands of district judges and makes the outcomes of toxic tort cases in federal courts turn on the prejudices of the particular judge rather than on principles of law. As Justice Stevens observed, "it bears emphasis that the Court has not held that it would have been an abuse of discretion to admit the expert testimony. The very point of today's holding is that the abuse of discretion standard of review applies whether the district judge has excluded or admitted the evidence."[78] Plainly, district judges are permitted to approve the weight-of-the-evidence methodology, if that is their disposition, and they are free to defer to the conclusions of the experts applying that method, if *that* is their disposition. District judges who take these two stances will admit the very testimony that was excluded in this case, and put the case to the jury. Appellate judges will be obliged to defer to the district court's exercise of discretion in this manner, just as *Joiner* held they were obliged to defer to an exercise of discretion in the opposite direction. Access to a jury trial in toxic torts cases will depend upon the unreviewable viewpoints of individual district court judges.

The question whether the weight-of-the-evidence methodology is a permissible methodology for determining whether general causation is "likelier than not" ought not to be left to district court discretion. As the methodology universally championed by federal agencies' scientists for making such

---

76. *Joiner*, 118 S. Ct. at 519.
77. *Id.* at 523 (Stevens, J., concurring in part and dissenting in part).
78. *Id.*

determinations, it has an imprimatur that should be proof against judicial rejection as "junk science."

Assuming, arguendo, that the *Joiner* opinion does not sanction such rejection, then it compounds the conundrum highlighted earlier in this paper. Empowering district judges to exclude expert testimony because they disagree with the sufficiency of the conclusions reached by the experts is ludicrous, unless one believes that district judges are better qualified than scientists to draw the inductive inferences that are the final step of the weight-of-the-evidence methodology. Obviously they are not. Thus, it seems even more evident that what is really going on beneath the surface is a transfer of the jury's "truth-determining" functions to the trial judge under the guise of an evidentiary ruling. What is more, the transfer is stacked against plaintiffs: if they persuade the trial judge, they earn the right to a jury's consideration. To win, they must win twice. If the trial judge is not persuaded, however, the case is over. The defendant need only win once.

III. SPECULATING ON THE FUTURE

What will happen in light of *Joiner?* Here are some preliminary speculations.

A. Neutral Experts

In his concurring opinion in *Joiner*, Justice Breyer offered district judges some advice "to help them overcome the inherent difficulty of making determinations about complicated scientific or otherwise technical evidence."[79] Among his suggestions were "the appointment of special masters and specially trained law clerks" (presumably, in both instances, those with scientific training), and the appointment of neutral experts as authorized in FRE 706.[80] Given the enormous discretion that *Joiner* vests in district judges, these suggestions surely have merit, but a danger exists.

Large portions of the scientific community are invested in one of various theories, which will make selection of the expert virtually the selection of an outcome. Plainly, courts are unlikely to appoint as neutrals those experts who regularly testify for one side or the other in toxic torts cases. However, a much larger population of scientists exists who are dependent upon the corporate world for substantial financing of their research activities, even though they are not ordinarily witnesses in court. These experts may have biases or conflicts that will influence their contributions. On the other side, a large population of scientists counsel and assist the government agencies that have aggressively pursued the weight-of-the-evidence methodology. They may have biases or conflicts cutting in the other direction.

If Justice Breyer's suggestion is to be followed in substantial volume by federal judges, it would be desirable to develop a corps of academic scientists who

79. *Id.* at 520 (Breyer, J., concurring).
80. *Id.* at 520–21.

make court-assistance their *only* extra-curricular activity, so that their evenhandedness is not compromised.

B. Forum Shopping

Plaintiffs will attempt assiduously to forum shop in order to keep their toxic tort cases out of federal court, at least in those states where the state courts do not apply a similar reliability screen for the admission of expert testimony. The manufacturers who are the obvious target of these suits will, most of the time, have citizenships diverse from the plaintiff. If they alone are sued, even in state court, the defendants will be able to remove to federal court and enjoy the benefit of the *Daubert-Joiner* evidentiary rules—rules that may, depending on the disposition of the federal district judge, have the practical effect of requiring plaintiffs to prove causation beyond a reasonable doubt. However, if even one person whose citizenship is the same as the plaintiffs is joined as a defendant, there will not be complete diversity and the case will not be removable.

This is not good news for doctors. Many cases involve a doctor who could be sued for alleged failure to diagnose the plaintiff's illness earlier or (in cases involving pharmaceuticals) for prescribing or not adequately warning about the side effects of the substance. To be sure, Rule 11 of the Federal Rules of Civil Procedure is a safeguard against frivolous joinder. That, however, is not meaningful protection for doctors. Often a non-frivolous claim can be advanced against a doctor, yet the plaintiff (absent concern about the admissibility of expert testimony) would opt not to join the doctor. This election might be based on respect and friendship, or it might be strategic: to secure a friendly witness on specific causation. *Daubert* and *Joiner* put increased pressure on the plaintiff to join the doctor to prevent removal to federal court.

Even when the case indisputably is destined for federal court, we can expect a heightened degree of forum shopping. Judges differ in their attitudes about the competing scientific views, and they are likely to be affirmed whichever approach they follow. So, plaintiffs' lawyers will seek out venues in which they are more likely to obtain judges who are "weight of the evidence" adherents.[81]

C. Inconsistent Judicial Rulings on Identical Issues

If courts of appeals understand *Joiner* to require them to stay their hands and affirm regardless of what the trial judge does, we can expect diametrically opposite rulings on the admissibility of expert testimony on the causative effects of substances that are the targets of multiple suits. These admissibility issues tend to be aired at the summary judgment stage—as in *Daubert* and *Joiner*—and it is not uncommon for the plaintiffs to introduce identical affidavits of the same experts in opposition to summary judgment in multiple cases. After all, the issue of general causation—whether substance *A* is capable of causing disease *B*—does not vary

81. As the conflicting views of the district and appellate courts in *Joiner* reflect, federal judges disagree about the proper approach.

from case to case, and it would be redundant to generate a new set of affidavits for each case.[82]

We can expect instances in which some federal judges rule that the plaintiff is not entitled to a trial on that issue, while other judges (perhaps in the same circuit, even in the same district) rule that the plaintiff *is* entitled to trial on the basis of exactly the same opinion from the same expert. The courts of appeals, if faithful to *Joiner*, will find themselves affirming both rulings. This phenomenon will be disquieting, if not promotive of a general disrespect within the populace for the integrity of federal judges. Of course, variance between courtrooms would be predictable even if these issues were given to juries in every case, but disagreement among juries is likely to be less offensive to shared communal values than sanctioned disagreement among judges.

In its *amicus* brief in *Joiner*, the United States, while arguing for an abuse of discretion standard of review, suggested that, with respect to recurring issues of the causative potential of a particular substance, the courts of appeals ought to have room to assure consistency over a range of cases.[83] Nothing in the Court's opinion in *Joiner* suggests a receptivity to that suggestion.

D. Appellate Disingenuity

In the wake of *Joiner*, it would not be astonishing to see some appellate court disingenuity in reviewing rulings admitting and excluding expert testimony. Seemingly deprived of the opportunity to achieve coherence through the articulation of doctrinal standards that elaborate upon *Daubert*, they may resort to simply declaring that a particular ruling is, on its facts, an abuse of discretion. Certainly, in the period before *Joiner*, judges on both sides of the philosophical divide were unhappy that issues of such magnitude would escape close appellate review. Third Circuit Judge Edward Becker advocated a "hard look" for rulings that excluded expert testimony and thereby non-suited plaintiffs in toxic torts cases,[84] while Fifth Circuit Judge Patrick Higginbotham championed review with "a sharp eye" of district court rulings that reflected a "'let it all in' philosophy" about such testimony.[85] "Abuse of discretion" is a standard with no well-defined meaning, and room exists for appellate work to be done below the radar-screen of likely Supreme Court review.

---

82. Variance will occur at the level of specific (as contrasted with general) causation, see *supra* note 50, but most of these cases (including *Daubert* and *Joiner*) are being resolved on the question of general causation.

83. Brief for the United States as Amicus Curiae in Support of Petitioners, General Electric Co. v. Joiner, 118 S. Ct. 512 (1997) (No. 96-188).

84. *In re* Paoli R.R. Yard PCB Litig., 35 F.3d 717, 750 (3d Cir. 1994).

85. *In re* Air Crash Disaster at New Orleans, Louisiana, 795 F.2d 1230, 1234 (5th Cir. 1986). *See also* Joy v. Bell Helicopter Textron, Inc., 999 F.2d 549, 569–70 (D.C. Cir. 1993).

E. Attention at the Stage of Senate Confirmation of Judicial Nominees

Recent years have witnessed heightened Senate scrutiny of judicial nominees with respect to their views on particular controversial issues such as abortions. Given that there are heavy lobbying interests on both sides of toxic tort litigation,[86] it is possible that a nominee's philosophical attitudes about scientific testimony will become a new litmus test for confirmation of district judges. After all, *Joiner* commits the fate of toxic tort litigation to the seemingly unreviewable discretion of district judges, so their appointment will likely shape the future of such litigation. If the track record of the contending behemoths on other issues before Congress is any indicator, this may spell even longer deadlocks over confirmation.

F. State Court Rejoinders

Finally, one or more state courts may become incensed at the federal undermining of state tort policy through the administration of federal evidence rules. If so, weapons are available to counter the federal incursion. States courts can alter their *substantive* tort rules to prevent the federal courts from non-suiting plaintiffs via the evidence rules. Here are a couple of examples.

Professor Margaret Berger, who served as Reporter for the Advisory Committee on the FRE that had this issue under consideration at the time *Daubert* was decided, has recently despaired that the problem can be solved coherently under a regime that makes causation the critical issue in toxic torts cases. She has advocated the removal of causation as a critical element in toxic torts cases, to be replaced by a measure of the defendant's assiduousness in researching the potential harmful consequences of its products.[87] This seems an extreme solution, but the prominence of its source suggests just how troublesome is the current state of play.

A less drastic solution might be to shift the burden of proof once plaintiffs have made a specified threshold showing of causation. For example, a state supreme court could prescribe that if plaintiffs showed a specified number of epidemiological studies in which there were elevated incidences of cancer, coupled with evidence that the substance is causing cancer in animal studies and/or evidence that the chemical composition of the substance is similar to other substances known to be carcinogenic, a presumption of causation would arise, which the defendant could overcome only by proving that the substance likelier than not *does not* cause such diseases. This approach would have the added credential that it would put the burden on the party with the best access to the pertinent data.[88]

---

86. The Association of Trial Lawyers of America (which includes most lawyers who specialize in representing plaintiffs in torts cases) and corporate America both lobby heavily on opposite sides of a variety of issues that come before Congress, including, for example, the federal products liability bill that has been pending for more than a decade.

87. Margaret A. Berger, *Eliminating General Causation: Notes Toward a New Theory of Justice and Toxic Torts*, 97 COLUM. L. REV. 2, 117 (1997).

88. *Id.*

The presumption approach would have two intriguing incentive effects. First, it would stimulate the greater corporate research that Professor Berger hopes to achieve, for companies would be required to show the safety of suspicious products. Second, it would have the amusing effect of reversing the present party-identification with competing scientific methodologies. American industry would at last discover the beauties of the weight-of-the-evidence methodology. Sadly, the plaintiffs' bar might forget them.